Development of a Malignancy Risk Classifier for Indeterminate Pulmonary Nodules Using Bronchoalveolar Lavage Immune Cell Signatures

Dear Lung Cancer Canada and GOMRG Committee:

This project aims to determine whether gene expression signatures of background lung immune cells collected from around lung nodules through bronchoalveolar lavage could be used to predict whether the nodule is likely cancer or not. If we can use signals from immune cells to predict the presence of nearby malignancy, then even otherwise non-diagnostic samples could become informative.

The classifier we propose to develop is expected to be readily clinically implementable, enabling earlier definitive management for lung cancer and reducing unnecessary procedures for patients with benign lesions.

Thank you for reviewing our application and for your support of lung cancer research. We sincerely hope to have the opportunity to pursue the work outlined.

Sincerely,

Julia Naso
Consultant Thoracic Pathologist, Vancouver General Hospital
Clinical Assistant Professor, Pathology and Laboratory Medicine, University of British Columbia

Summary of Proposed Research

Pulmonary nodules are commonly detected on CT scans incidentally or through lung cancer screening programs. However, benign nodules (e.g. scars, granulomas and organizing pneumonia) may be indistinguishable from lung cancer on imaging, particularly without interval follow-up. Patients commonly wait through years of imaging follow-up before diagnosis, and may have unnecessary major surgery: up to 48% of patients undergoing surgical lung resection do not have prior biopsy diagnosis,[1,2] and 10-14% of such 'undiagnosed' nodules are found after surgery to be benign.[1,3,4] These unnecessary surgeries have considerable risks, with lobectomy for moderate-risk patients having 3% mortality and 21% risk of cardiopulmonary complications.[5] Conversely, delaying surgery to allow longer imaging follow-up may allow even small tumors to seed incurable metastases. Early metastatic divergence is predicted to occur when tumors are less than 8 mm,[6] and even the earliest stage lung cancer (stage IA, measuring ≤ 1 cm) has a ~15% chance of recurring as metastatic disease following clinically complete resection.[7]

Current attempts at diagnosis on biopsy/cytology have very limited sensitivity because they rely on malignant cells being seen. If we can use signals from background reactive cells to predict the presence of nearby malignancy, then even otherwise non-diagnostic samples could become informative. Bronchoalveolar lavages (BALs) contain cells rinsed out of a lung region through the airways, and are commonly performed for microorganism cultures (to exclude active infection) and screening for malignant cells. BALs capture malignant cells in less than a quarter of patients with lung cancer[8–10] but almost always show abundant immune cells (Figure 1A). Immune cells in the lung cancer microenvironment have signatures distinct from those in normal lung.[11–14] These differences are observable in BAL specimens,[15,16] and we hypothesize such differences would also be present in comparison with benign nodules. The development of a malignancy risk classifier requiring only the non-neoplastic cells in BALs would enable these minimally invasive samples to be informative for a much greater proportion of patients.

Our **overall objective** is to identify non-neoplastic cell signatures in BAL specimens from lungs with benign or malignant nodules. Our central **hypothesis** is that the expression of immune signaling genes in BAL specimens can be used to accurately distinguish benign and malignant nodules. Our **approach** will be to use NanoString Immune Panel Assays on archival BAL specimens (Figures 1B and 2). In **preliminary data** (n=24) we identify alpha-beta T cell activation as downregulated and MHC class II assembly as upregulated in false negative BALs from lung adenocarcinoma patients compared to BALs from patients with benign lung nodules (q-values<0.05, Figure 3). A gene expression signature of *ICAM2*, *C3*, *CARD9*, *TGFB2*, *PNMA1*, and *IL22RA2* was able to discriminate adenocarcinoma cases from benign controls with 100% sensitivity and specificity (Figure 4A,B). This finding was reproduced with 100% sensitivity and specificity using an external gene expression microarray dataset of BAL specimens (malignant n=9; benign n=6).[15] NanoString data for an additional 24 cases is currently being generated, which will exhaust UBC New Faculty Award funds for this project. The requested funding is essential to bring the total sample size up to 96, necessary to have suitable statistical power (0.86) for detecting differences between groups with a false discovery rate (FDR) of 0.05, based on the mean and standard deviation for the 50 most differentially expressed genes in the first 24 cases. Profiling this well-powered cohort is critical to avoid overfitting in classifier generation, ensuring results are generalizable to future validation cohorts.

The **rationale** for this approach is that tumor cells are often not present in BAL samples, whereas immune cells are abundant, and the current clinical practice of disregarding immune cells in BAL specimens results in missed opportunities for early diagnosis. Compared to other

bronchoscopy specimens, BALs enable greatest sampling of immune cells in alveolar parenchyma/airspaces with minimal upper airway contamination. The applicability of our classifier to specimens prepared in existing clinical workflows and able to be stored while conventional diagnostic investigations are completed supports rapid clinical implementation. Upon completion, we expect to have developed a robust and clinically feasible malignancy-risk classifier.

Specific Aims: Aim 1: Perform gene expression profiling of BAL specimens: Archival clinical BAL cell blocks from an additional 48 patients (for a total cohort size of 96, half with lung carcinoma and half with benign lung nodules), will be profiled using the NanoString PanCancer Immune Profiling nCounter Assay (770 genes, suitable for as little as 25 ng RNA). BAL specimens that were culture positive for pathogenic organisms or had definite malignant cells identified on clinical cytomorphology assessment will be excluded as these specimens are already sufficient for clinical diagnosis. Including only the remaining specimens ensures that the identified signatures are applicable to specimens for which additional diagnostic tools are needed. To ensure that the benign and malignant group designations are correct, all specimens will have either diagnostic confirmation on another specimen or 2 years of imaging follow-up to exclude malignancy. Acquisition of the proposed sample size is feasible based on ~250 BALs having been performed each year at our site and blocks up to 10 years old being suitable for NanoString assessment (probe-based assay, robust to degradation). Malignant and benign cohorts will be balanced for never versus former smoking, autoimmune disease, sex, and time since specimen collection. Currently smoking patients will be excluded to avoid confounding effects of smoking-related immune dysregulation; possible generalizability of the classifier to currently smoking patients will be explored in subsequent validation studies.

Aim 2: Identify malignancy-risk prediction signatures: Differential gene expression and pathway activity in benign versus malignant samples will be identified. We will use nSolver for quality assessment, normalization and differential expression analysis with false discovery rate correction. Gene set enrichment analysis will use the R clusterProfiler package, and the glmnet package will be used for penalized Lasso regression models. A neural network classifier for malignancy prediction will be implemented in PyTorch with SHAP score assessment for feature selection.[17,18] Classifier performance will be assessed in five-fold cross validation. The genes with greatest importance in the classifier will be identified, and a final gene set of a size practical for clinical assessment in custom NanoString assays (up to ~40 genes) will be determined.

Aim 3: Comparison to external datasets and clinical/histologic features: To validate differential expression, identify the cell types with differential gene expression, and assess the generalizability of our findings we will assess genes of high importance to the classifier in publicly accessible single cell gene expression profiling data[19–22] and public BAL gene expression datasets[15,16] from lungs with and without cancer (example in Figure 4C,D). We will also assess our BAL gene expression data for associations with other clinical and histologic features of our study patients and their tumors, to understand features potentially impacting or impacted by immune microenvironments. The final classifier gene set will be refined based on these data, to produce a malignancy risk classifier likely to have robust performance in future prospective studies using custom NanoString probe sets.

UBC **Research Ethics Board approval** with a waiver of patient consent has been obtained. Our **team** of clinician scientist pathologists, cancer genome researchers, lung cancer immunologists and bioinformaticians is expertly positioned to deliver results on this project.

References:

1. Onwugbufor MT, Soni ML, Predina JD, et al. Lobectomy for Suspected Lung Cancer Without Prior Diagnosis. *Ann Thorac Surg*. 2023;116(4):694-701. doi:10.1016/j.athoracsur.2023.05.017

2. Sihoe ADL, Hiranandani R, Wong H, Yeung ESL. Operating on a suspicious lung mass without a preoperative tissue diagnosis: pros and cons. *Eur J Cardio-Thorac Surg Off J Eur Assoc Cardio-Thorac Surg*. 2013;44(2):231-237; discussion 237. doi:10.1093/ejcts/ezs671

3. Archer JM, Mendoza DP, Hung YP, Lanuti M, Digumarthy SR. Surgical Resection of Benign Nodules in Lung Cancer Screening: Incidence and Features. *JTO Clin Res Rep*. 2023;4(12):100605. doi:10.1016/j.jtocrr.2023.100605

4. Giles AE, Teferi Y, Kidane B, et al. Lung Resection Without Tissue Diagnosis: A Pragmatic Perspective on the Indeterminate Pulmonary Nodule. *Clin Lung Cancer*. 2021;22(5):e774-e781. doi:10.1016/j.cllc.2021.02.011

5. Gooseman MR, Falcoz PE, Decaluwe H, Szanto Z, Brunelli A. Morbidity and mortality of lung resection candidates defined by the American College of Chest Physicians as “moderate risk”: an analysis from the European Society of Thoracic Surgeons database. *Eur J Cardio-Thorac Surg Off J Eur Assoc Cardio-Thorac Surg*. 2021;60(1):91-97. doi:10.1093/ejcts/ezab028

6. Al Bakir M, Huebner A, Martínez-Ruiz C, et al. The evolution of non-small cell lung cancer metastases in TRACERx. *Nature*. 2023;616(7957):534-542. doi:10.1038/s41586-023-05729-x

7. Carr SR, Wang H, Hudlikar R, et al. A Unique Gene Signature Predicting Recurrence Free Survival in Stage IA Lung Adenocarcinoma. *J Thorac Cardiovasc Surg*. 2023;165(4):1554-1564. doi:10.1016/j.jtcvs.2022.09.028

8. Bezel P, Tischler V, Robinson C, et al. Diagnostic Value of Bronchoalveolar Lavage for Diagnosis of Suspected Peripheral Lung Cancer. *Clin Lung Cancer*. 2016;17(5):e151-e156. doi:10.1016/j.cllc.2015.12.012

9. Nguyen-Dang K, Bui-Thi HD, Duong-Minh N, et al. The Role and Associated Factors of Liquid-Based Cytology of Bronchoalveolar Lavage Fluid in Lung Cancer Diagnosis: A Prospective Study. *Cureus*. 2023;15(11):e48483. doi:10.7759/cureus.48483

10. Sharma A, Naso J. The diagnostic utility of bronchoalveolar lavage in the evaluation of pulmonary lesions: A two-year retrospective analysis. Published online 2025.

11. Taniguchi S, Matsui T, Kimura K, et al. In vivo induction of activin A-producing alveolar macrophages supports the progression of lung cell carcinoma. *Nat Commun*. 2023;14(1):143. doi:10.1038/s41467-022-35701-8

12. Casanova-Acebes M, Dalla E, Leader AM, et al. Tissue-resident macrophages provide a pro-tumorigenic niche to early NSCLC cells. *Nature*. 2021;595(7868):578-584. doi:10.1038/s41586-021-03651-8

13. Noy R, Pollard JW. Tumor-associated macrophages: from mechanisms to therapy. *Immunity*. 2014;41(1):49-61. doi:10.1016/j.immuni.2014.06.010

14. Chang CY, Armstrong D, Corry DB, Kheradmand F. Alveolar macrophages in lung cancer: opportunities challenges. *Front Immunol*. 2023;14:1268939. doi:10.3389/fimmu.2023.1268939

15. Kuo CHS, Liu CY, Pavlidis S, et al. Unique Immune Gene Expression Patterns in Bronchoalveolar Lavage and Tumor Adjacent Non-Neoplastic Lung Tissue in Non-Small Cell Lung Cancer. *Front Immunol*. 2018;9:232. doi:10.3389/fimmu.2018.00232

16. Kalkanis A, Papadopoulos D, Testelmans D, Kopitopoulou A, Boeykens E, Wauters E. Bronchoalveolar Lavage Fluid-Isolated Biomarkers for the Diagnostic and Prognostic Assessment of Lung Cancer. *Diagn Basel Switz*. 2022;12(12):2949. doi:10.3390/diagnostics12122949

17. Ferrato MH, Marsh AG, Franke KR, et al. Machine learning classifier approaches for predicting response to RTK-type-III inhibitors demonstrate high accuracy using transcriptomic signatures and ex vivo data. *Bioinforma Adv*. 2023;3(1):vbad034. doi:10.1093/bioadv/vbad034

18. Zhang A, Akbari V, Galbraith A, et al. Cross-platform methylation-based site of origin classification for squamous cell carcinomas. *Mod Pathol*. 2025;in press, journal pre-proof.

19. Yanagawa J, Tran LM, Salehi-Rad R, et al. Single-Cell Characterization of Pulmonary Nodules Implicates Suppression of Immunosurveillance across Early Stages of Lung Adenocarcinoma. *Cancer Res*. 2023;83(19):3305-3319. doi:10.1158/0008-5472.CAN-23-0128

20. Renaut S, Saavedra Armero V, Boudreau DK, et al. Single-cell and single-nucleus RNA-sequencing from paired normal-adenocarcinoma lung samples provide both common and discordant biological insights. *PLoS Genet*. 2024;20(5):e1011301. doi:10.1371/journal.pgen.1011301

21. Kim N, Kim HK, Lee K, et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat Commun*. 2020;11(1):2285. doi:10.1038/s41467-020-16164-1

22. Bischoff P, Trinks A, Obermayer B, et al. Single-cell RNA sequencing reveals distinct tumor microenvironmental patterns in lung adenocarcinoma. *Oncogene*. 2021;40(50):6748-6758. doi:10.1038/s41388-021-02054-3

23. Gould MK, Tang T, Liu ILA, et al. Recent Trends in the Identification of Incidental Pulmonary Nodules. *Am J Respir Crit Care Med*. 2015;192(10):1208-1214. doi:10.1164/rccm.201505-0990OC

24. Canadian Medical Imaging Inventory 2022–2023: CT: CMII Report [Internet]. Ottawa (ON): Canadian Agency for Drugs and Technologies in Health; 2024 Aug. Available from: https://www.ncbi.nlm.nih.gov/books/NBK607343/.

25. Canadian Partnership Against Cancer. Expanding Lung Cancer Screening. https://www.partnershipagainstcancer.ca/annual-report/2024-2025-annual-report/year-in-review/early-accurate-diagnosis/expanding-lung-cancer-screening/#:~:text=Lung%20cancer%20remains%20the%20leading,saving%204%2C700%20to%205%2C800%20lives.

Appendix:

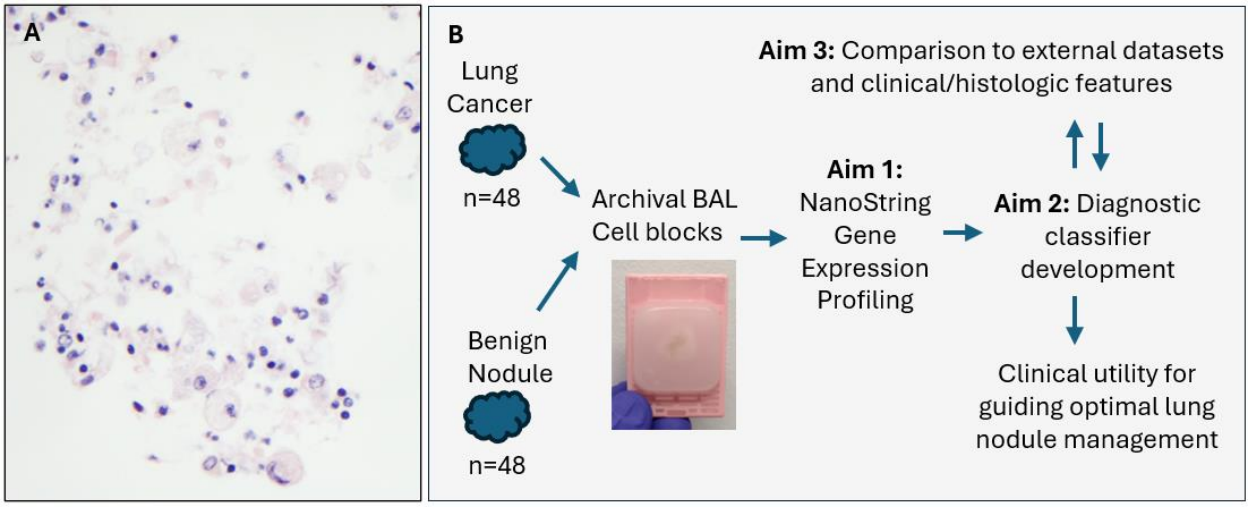

Figure 1: (A) Bronchoalveolar lavage cell blocks (hematoxylin & eosin stain, 400X) typically contain macrophages, lymphocytes and bronchial epithelial cells in variable proportions, with occasional neutrophils present. Understanding the precise cell types present and their activities requires molecular studies. **(B)** Approach to project objectives.

Activity	**2026**			
	Q1	**Q2**	**Q3**	**Q4**
Specimen identification and retrieval				
RNA extractions and NanoString assays				
Differential gene expression analysis, classifier development				
Comparison with single cell and external data				
Manuscript preparation				

Figure 2: Project timeline

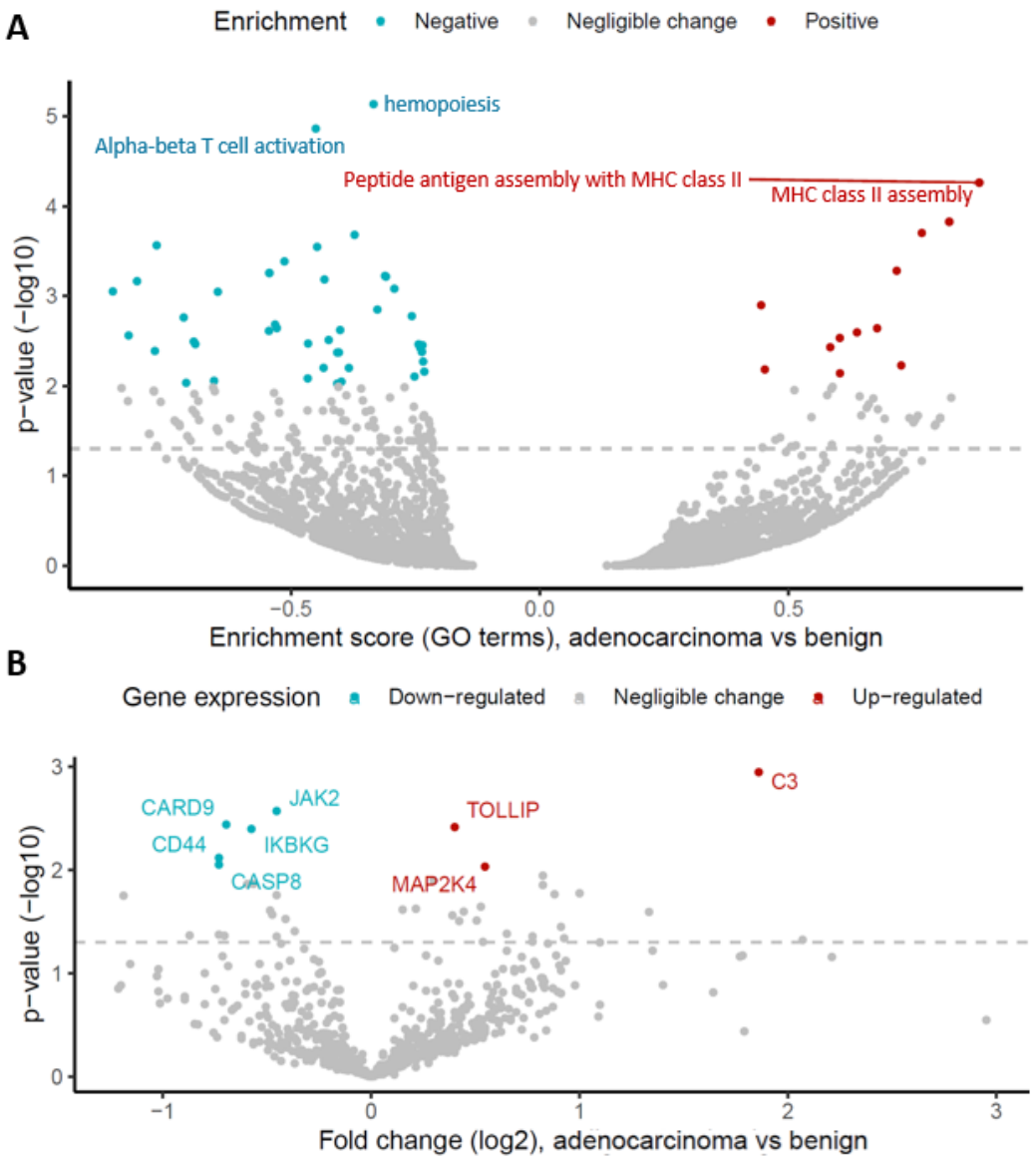

Figure 3: BAL NanoString profiling using NanoString Immune Panel Assays shows differential activity of (A) gene ontology groups and (B) genes between cytologically negative samples from patients with benign (n=12) and adenocarcinoma (n=12) lung nodules. Colored points indicate statistically significance (FDR < 0.05) after multiple testing correction.

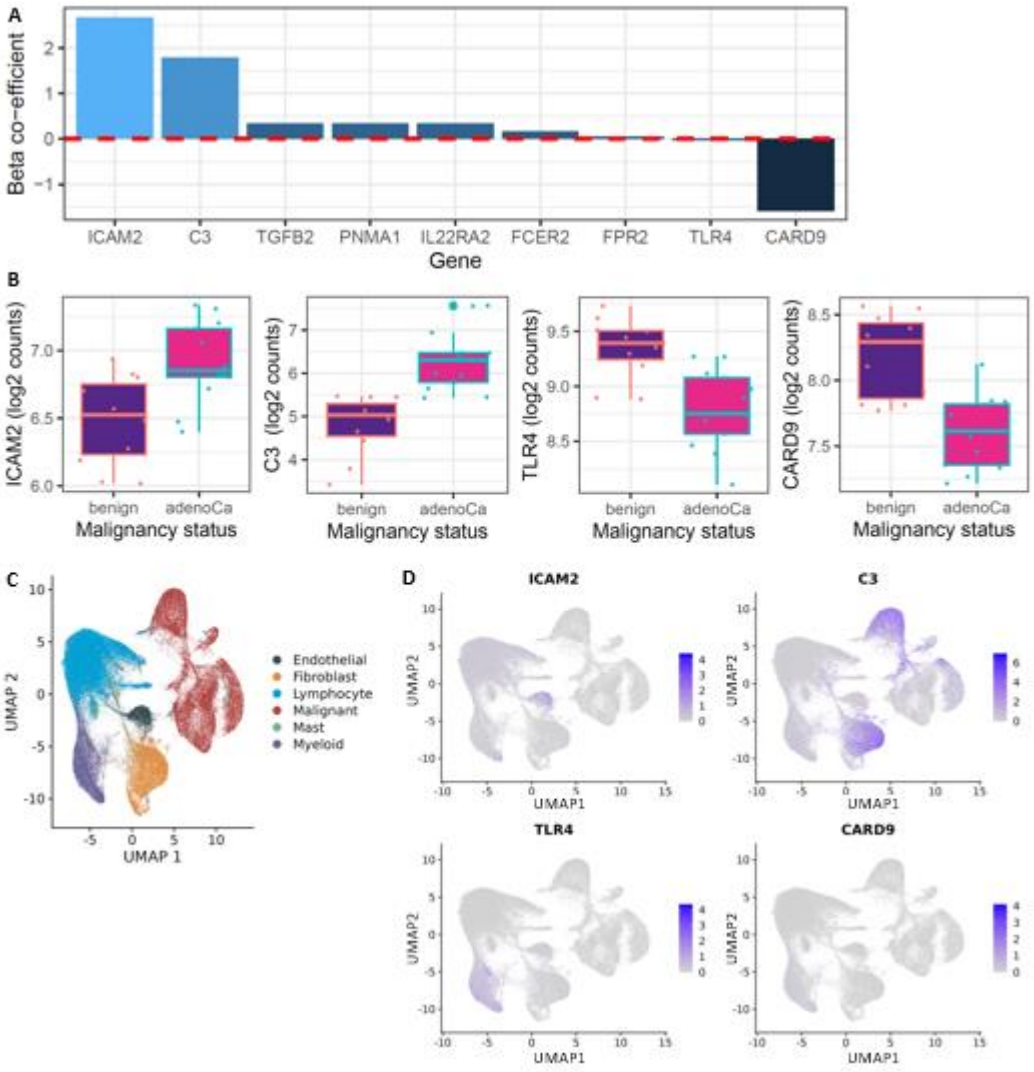

Figure 4: (A) Lasso regression with five-fold cross-validation identified genes predictive of adenocarcinoma vs. benign lung nodules. (B) Expression of key predictive genes in BALs from patients with adenocarcinoma vs benign nodules. (C) Single cell gene expression profiling of lung carcinoma resection specimens enables clustering of cells by type and (D) can be used to identify the cell types expressing the four key predictive genes. Moderate expression in lymphocyte or myeloid cell clusters is seen for all four genes, supporting the notion that the differential expression signals could be from different populations/activities of immune cells.

Impact Statement:

Bridging the gap between CT screening and early diagnosis for lung cancer requires an ability to distinguish whether nodules detected on imaging are benign or malignant. Each year there are ~1.9 million Canadians with an incidental lung nodule identified each year[23,24] and thousands more with screen-detected pulmonary nodules.[25] These patients are in critical need of improved minimally invasive strategies for determining whether their lung nodules are likely malignant.

Upon completion of this project we expect to have developed an accurate and clinically feasible classifier for malignancy-risk, based on assessment of routinely collected minimally invasive samples. The classifier we propose would be the first capable of using otherwise uninformative BAL cell blocks to risk-stratify indeterminate lung nodules. Prior studies of BALs have shown differences in gene expression between samples from lung cancer patients and normal lungs (i.e. patients without lung nodules), and this study fills an urgent need to adapt such signatures to lung cancer vs benign nodule discrimination, the scenario seen in clinical practice.

We expect our malignancy-risk classifier to provide critical information for guiding decisions on surveillance intensity, biopsy and surgical resection for patients with screen-detected or incidental lung nodules, ultimately enabling lung cancer patients to receive more timely diagnoses and more effective lung cancer management, while also reducing unnecessary procedures for patients with benign lesions.

Our approach is designed for rapid clinical translation through (i) our leveraging of archival tissues with pre-existing long-term follow-up data, (ii) our use of an assay type that can feasibly be performed in clinical laboratories at low cost (clinical NanoString assays are performed weekly at our hospital, with reagent costs <$100/sample) and (iii) our use of a specimen type that is already routinely collected on patients with lung nodules and prepared in the clinically standard fashion. The compatibility of our classifier with existing clinical workflows makes it well suited to rapid clinical implementation.

The present study provides the requisite scientific basis for prospective test validation and clinical trials, which we will pursue following completion of this study. Local prospective implementation of the classifier through a custom NanoString panel applied to BALs as a laboratory-developed test would be supported through Dr. Naso's clinical practice in molecular and thoracic pathology, enabling acquisition of data for wider health regulator clearance. Overall our team is well-positioned to deliver scientifically robust findings readily translatable into improved care for lung cancer patients in Canada and beyond.

Public Summary

Lung nodules may be detected on CT scans performed to detect lung cancer in patients at risk (screening), or may be unexpected findings on imaging for other concerns. Such nodules may or may not be lung cancer. Distinguishing these possibilities on initial CT scans is challenging as lung cancer and scars may appear similar. Patients may wait through multiple years of imaging follow-up to help predict whether a nodule is cancer, and may have complications from biopsies, which often do not provide a diagnosis as cancer cells are often not picked up. Many surgeries to remove a lung nodule are done without prior definite diagnosis, and many of these surgeries are later found to be unnecessary as the nodule turns out to be benign.

Safely and promptly determining whether a lung nodule is likely to be cancerous is thus critical for allowing high-risk nodules to proceed to management (e.g. through sooner interval follow-up or proceeding to repeat biopsy or surgery), and for patients with low-risk nodules to avoid unnecessary procedures and have less intensive surveillance. Current attempts at diagnosis often are not informative because they rely on malignant cells being seen. If we can use signals from immune cells collected from the lungs to predict the presence of nearby cancer, then even otherwise unhelpful samples could become informative. Immune system cells in the lung can be rinsed out and collected through the airways in a minimally invasive procedure, bronchoalveolar lavage (BAL). BALs are already commonly performed on patients with lung nodules. This project aims to determine whether immune cells collected from around a lung nodule through BAL could be used to predict whether the nodule is likely to be a cancer, to guide next steps in management.

Initial data from our team has identified different genes ‘turned on’ in BAL samples from patients with versus without lung cancer. The samples that these signals were found in had no identifiable cancer cells, such that the differences in gene activity appear to be from different populations or activities of non-cancer cells. The genes with different activity had immune-related functions and expression in immune cell types, and thus may reflect immune system responses that are specific to lung cancer. These signals may therefore able be to inform on whether lung cancer is likely to be present.

We propose to develop a cancer-risk classifier that uses leftover cells from BAL specimens. The requested funding will enable us to profile enough BAL samples to be confident that the classifier we develop will have similar performance on future datasets, including in future possible clinical use. As our classifier could be applied to existing workflows at low cost, we anticipate rapid clinical implementation, enabling earlier management for lung cancer and reducing unnecessary procedures for patients with benign lesions. Such improvements in patient care have the potential to benefit the ~1.9 million Canadians with incidental lung nodules identified each year and the thousands more with screen-detected pulmonary nodules.

Budget

Block and slide retrieval: $26/case x 48 cases = ~**$1,250**

RNA extraction: $50/case x 48 cases **= $2,400**

NanoString immune panel assay: $444 per assay x 48 samples: ~**$21,350**

The cost per assay includes $394 for the probeset, $38 for the mastermix, and shipping costs based on our prior orders.

VGH Molecular Laboratory Scientist Georgina Barnabas (salary separately funded) will conduct the NanoString assays, and data will be analysed by pathology resident Mark Trinder.

Total request: **$25,000**

List of Investigators:

Principal Investigator: Dr. Julia Naso, MD/PhD, FRCPC
Thoracic and Molecular Pathologist, Vancouver General Hospital
Clinical Assistant Professor, University of British Columbia
Molecular and Advanced Pathology Core (MAPcore) Co-Scientific Lead
Medical Research Director for Anatomical Pathology, Vancouver General Hospital

Co-Investigator: Dr. Mark Trinder, MD/PhD
Diagnostic and Molecular Pathology Resident
University of British Columbia

Co-Investigator: Dr. Marco A. Marra, PhD, FRS(C), FCAHS, O.B.C., O.C.
Distinguished Scientist, Michael Smith Genome Sciences Centre, BC Cancer Research Institute
UBC Killam Professor, Department of Medical Genetics
Tier 1 Canada Research Chair, Genome Science

Co-Investigator: Cathy Yan
PhD Student, Department of Medical Genetics, University of British Columbia
Canada's Michael Smith Genome Sciences Centre, BC Cancer Research Institute

Co-Investigator: Dr. Katey S.S. Enfield, PhD
Scientist, Integrative Oncology, BC Cancer Research Centre
Assistant Professor, Department of Pathology and Laboratory Medicine, University of British Columbia

Co-Investigator: Dr. Emilia Lim, PhD
Assistant Professor, Biochemistry and Molecular Biology, University of British Columbia
Investigator, Edwin S.H. Leong Centre for Healthy Aging

Pathology and Laboratory Medicine
Regional Administration
JPN 1, Room 1250
855 W. 12th Avenue
Vancouver, BC V5Z 1M9
T: 604-875-4577 F:604-875-5707

Sept 9th, 2025

Re: Geoffrey Ogram Memorial Research Grant Application of Dr Julia Naso

Dear Lung Cancer Canada Team:

This letter is in support of the application of Dr. Julia Naso, a consultant Anatomical Pathologist at VGH and UBC Clinical Assistant Professor. I confirm that her proposed project is feasible within our institution and its collaborating research centers.

We are closely affiliated with the Molecular and Advanced Pathology Core (MAPcore), a clinically accredited core facility which collaborates with the clinical molecular laboratory to run NanoString assays. Dr. Naso herself has 30% protected time for research. As a teaching center for residents and medical students, our department also supports and encourages trainee involvement in research projects such as this.

I am confident that our institution and affiliated centers can provide the resources needed to support Dr. Naso's proposed project.

Sincerely,

Dr. David Schaeffer
Department Head VCH Pathology and Laboratory Medicine